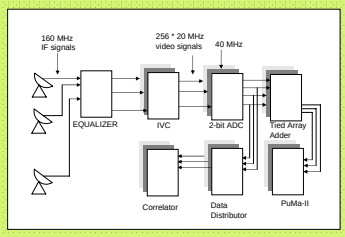
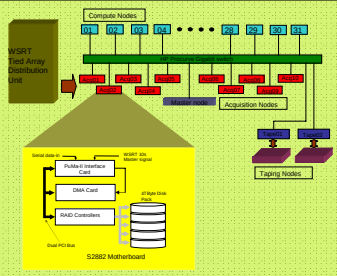
Software Aspects of PuMa-II

Ramesh Karuppusamy[1,2], Willem van Straten[3], Ben Stappers[1,2]
[1]Stichting ASTRON, Dwingeloo, The Netherlands
[2] Sterrenkundig Instituut "Anton Pannekoek", University of Amsterdam, The Netherlands
[3]University of Texas, Brownsville, USA.

PuMa-II is the second generation Pulsar Machine installed at the Westerbork Synthesis Radio Telescope (WSRT) and is open to the Astronomy Community since December 2005. The instrument can process the complete 160 MHz bandwidth offered by WSRT, as eight 20 MHz subbands, offering a routine time resolution of 25 ns and the possibility of ~3 ns when the subbands are combined. Up to 14 telescopes of WSRT can be phase coherently added in the Tied Array Adder Unit, giving sensitivity of a ~93-m aperture.

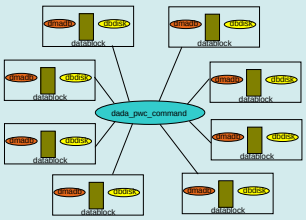




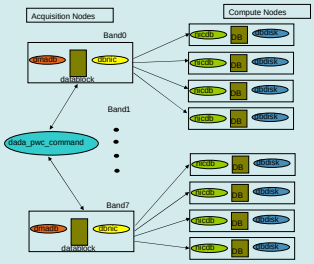
PuMa-II hardware is composed of a 42-node Linux Cluster. The cluster is logically separated into acquisition and compute nodes. See diagram above. Each node consists of dual Opertron processors clocked at 2GHz, a high performance mother board and 2GBytes of physical RAM. The complete instrument features ~400 GFlops of computational power, ~50 TeraBytes of storage space and can sustain a data throughput of 640 Mbytes/sec for 12 hours.

Local Recording Mode

PuMa-II in Local Acquisition Mode. The raw data is dumped from the data block (a 800 MB shared memory segment), into a 4 TByte local RAID disk pack. The central control, dada_pwc_cmd is the coordinator and the interface to the outside world. Each node sustains 80 Mbytes/sec sustained record rate. The "dmadb" is the direct memory access client, and "dbdisk" is the disk-write client.



Distributed Recording/Processing Mode



PuMA-II in Distributed Recording Mode. The raw data is pumped into the network interface from the data block of acquisition nodes. Each of this stream is targeted to at least 4 compute nodes to sustain 80 Mbytes/sec throughput. Each compute node fill it's data block. Depending on the pulsar period and dispersion measure, data is processed immediately offering real time processing or flushed to the disk. Each node has 500 GByte of storage space. "dbnic" and "nicdb" are the network interface clients. "dbdisk" can be replaced by "dspsr" to perform real-time processing.

Pulsar Signal Analysis Basics:

- ❑ Pulsar signals can be seen as amplitude modulated noise signals.
- ❑ Due to the propagation effects, pulsar signals suffer dispersion and scattering in the interstellar medium. This can completely smear out the pulsar signal, if left uncorrected.
- ❑ Dispersion can be undone before or after detection and is called coherent or incoherent dedispersion respectively. Coherent dedispersion needs raw voltages from the telescope.
- ❑ Coherent dedispersion corrects dispersion completely and incoherent dedispersion corrects partially, and hence effective time resolution can be affected.

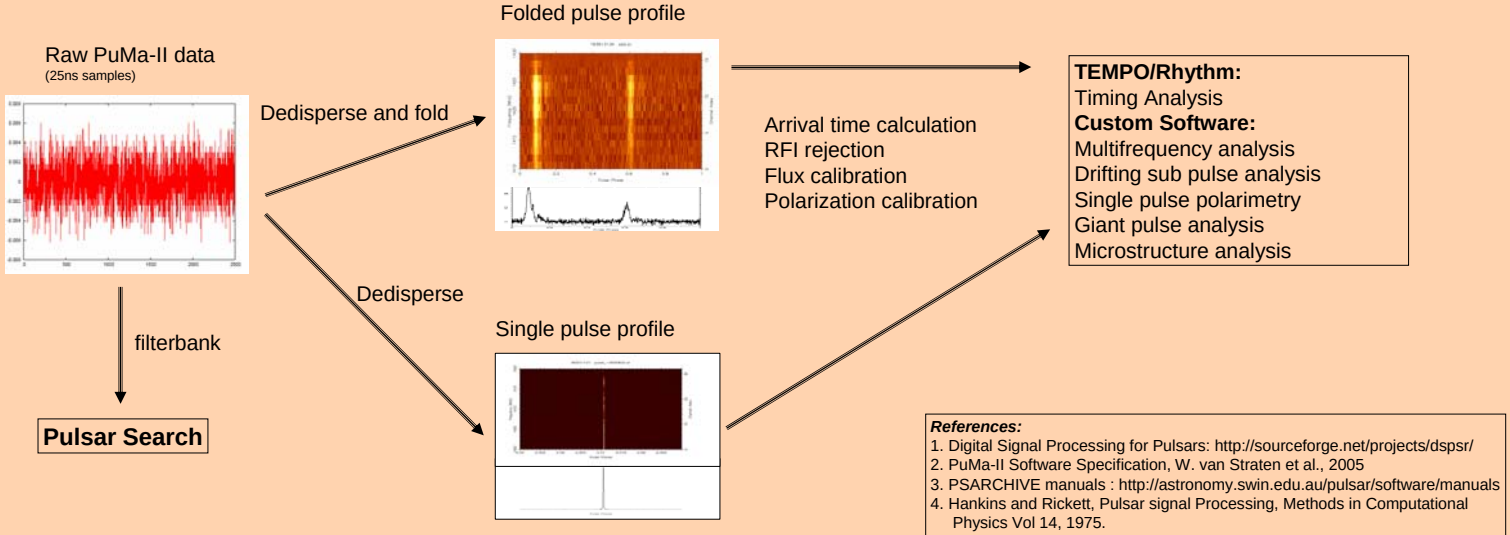
Acquisition Software:

PuMa-II software acquires data locally or in a distributed mode. In local mode digitized raw voltages of the two orthogonal polarizations are dumped to local disks, and in distributed recording mode, this data is dumped to remote disks. The software :

- ❑ is modular, scalable, multi-threaded C-Code
- ❑ uses EDT DMA Library routines for high speed acquisition.
- ❑ makes use of Linux device driver to control PuMa-II Interface Card to synchronize and time stamp data.
- ❑ manages high data throughput using a three stage buffering scheme – a hardware based FIFO at the lowest level, a 50ms kernel level buffer followed by 10-second buffer at user space.
- ❑ Uses sockets, pthreads, shared memory and semaphores
- ❑ runs off a Linux-based cluster environment based on Linux kernel version 2.6.14.
- ❑ provides a simple slave-like interface of PuMa-II to the outside world
- ❑ consists of the three modules – a direct memory access (DMA) client, and a disk write client, network interface client.
- ❑ can be used for real-time processing mode, if data is sent to processing software in compute nodes directly.

Processing Software:

PuMa-II uses DSP for Pulsars[1] (dspsr) to process raw data. The raw voltages are coherently dedispersed or filterbanked using dspsr. Final data products are single pulses or folded pulse profiles in PSRFITS format and contains all four stoke's parameters. The PSRCHIVE suite is used to analyze pulse arrival times, flux calibration, polarimetric models and RFI rejection on the data products. Other utilities exist to produce pulsar archive data plots, modify archival information, estimate arrival time and to study single pulses. For more information, see [3]



References:
1. Digital Signal Processing for Pulsars: http://sourceforge.net/projects/dspsr/
2. PuMa-II Software Specification, W. van Straten et al., 2005
3. PSARCHIVE manuals : http://astronomy.swin.edu.au/pulsar/software/manuals
4. Hankins and Rickett, Pulsar signal Processing, Methods in Computational Physics Vol 14, 1975.
5. EVNFRA Data Distributor Unit, NFRA Note 636, Rob Millenaar, Dec 1997.